Structural principles in cell-cycle control: beyond the CDKs

Jane A Endicott and Martin EM Noble*

Retinoblastoma protein (Rb) interacts with cyclin-dependent kinases and regulates the transcription of genes necessary for progression through the S phase of the cell cycle. Clues to the atomic mechanisms involved are offered by the structure of the two pocket regions of Rb in complex with a short peptide from a viral oncoprotein. Structures of cyclins, Rb and TFIIB reveal that a common motif occurs in proteins regulating three consecutive events of cell-cycle control.

Address: Laboratory of Molecular Biophysics, The Rex Richards Building, South Parks Road, Oxford OX1 3QU, UK.

*Corresponding author.
E-mail: martin@biop.ox.ac.uk

Structure 15 May 1998, 6:535–541
http://biomednet.com/elecref/0969212600600535

Introduction

The *Rb* gene was cloned and characterised through the correlation of genetic lesions at its locus with the development of retinal tumours in children [1–3]. Since then it has been shown that the genomes of cells from a wide variety of tumour types, notably small-cell lung cancer, osteosarcomas and bladder carcinomas, contain mutations that inactivate Rb [4]. In the cells of many other types of tumours, Rb function is compromised through mutations in molecules that regulate Rb activity such as cyclin D and the cyclin-dependent kinase (CDK) inhibitor p16 (reviewed in [5]). Together, this evidence has identified Rb as a tumour suppressor, able to inhibit cell division through its unique position at the molecular interchange between the processes that coordinate cell-cycle progression and the downstream events that initiate the transcription of genes required for the G1→S transition and DNA replication. The region of Rb required for growth inhibition includes two regions conserved between Rb family members, together termed the A–B pocket. Following on from the determination last year of the structure of the Rb A pocket region by Kim and Cho [6], Lee *et al.* have now described the structure of this A–B pocket in complex with a peptide ligand [7].

The role of Rb

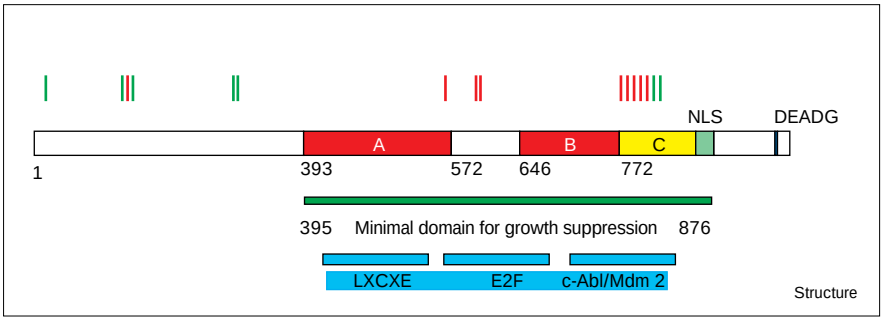
Rb negatively regulates cell-cycle progression largely by repressing the activity of the transcription factor E2F [5]. Elevation of E2F activity has been shown to be sufficient to drive quiescent cells into S phase [8]. Binding of hypophosphorylated Rb to E2F down regulates its transactivation activity and, at certain promoter sequences, Rb and E2F can form a complex that actively represses transcription [9]. E2F transcription factors are heterodimers composed of an E2F subunit and a DP subunit. Five E2F and two DP family members have been identified (reviewed in [10]). Rb characteristically, but not exclusively, associates with E2Fs containing the E2F-1, -2 or -3 subunits, whereas the Rb-related proteins p107 and p130 preferentially associate with E2F-4 and E2F-5 (reviewed in [5,11]). Active E2F is released following hyperphosphorylation of Rb by members of the CDK family. There are sixteen CDK consensus phosphorylation sites in Rb and early mapping experiments suggested that many of them are functional *in vivo* [12–14] (Figure 1). Cyclin D1-CDK4 and cyclin A/E-CDK2 phosphorylate different Rb sites *in vivo* [15,16]. Distinct subsets of Rb phosphorylation sites modulate Rb binding to E2F, to LXCXE motif-containing proteins and to c-Abl [16,17]; reviewed in [18].

The minimal Rb sequence required for its growth suppression activities has been mapped to the C-terminal region of the protein (residues 395–876). This sequence contains three distinct protein-binding sites, through which it associates with a number of proteins in addition to E2F [19]. The A–B pocket consists of two distinct sequences that are conserved between pocket-protein family members, p107, p130 and Rb. The pocket region of Rb in isolation is able to bind oncoproteins of a number of DNA tumour viruses, notably E1A from adenovirus, the large T antigen from SV40 and E7 from human papillomavirus (HPV) [20]. These proteins contain a characteristic LXCXE motif. The same motif is found in a number of other proteins, many of which are transcription factors that are able to bind to the A–B pocket region of Rb [14].

Although E2F does not have the LXCXE motif, binding of E2F to Rb requires the A–B pocket region, as well as additional residues towards the C terminus of the C domain (sometimes collectively referred to as the large pocket). The C domain also forms a binding site for c-Abl tyrosine kinase and the transcription factor Mdm2 [21,22]. Rb can simultaneously bind various combinations of these proteins, an observation which provides part of the evidence for the current model of Rb's function as a 'molecular matchmaker' in the assembly of macromolecular complexes [19].

The ability of Rb to bind to multiple transcription factors may in part explain its function as a specific repressor of certain genes that are transcribed by polymerase II (pol II)

Figure 1



Summary of the binding functions and phosphorylation sites in Rb [19]. The A-, B- and C-pocket regions of Rb are marked. The mapped binding sites for the LXCXE motif, E2F and the c-Abl tyrosine kinase are indicated with bars. The nuclear localisation signal (NLS) and the DEADG motif are also shown. Serine (249, 567, 608, 612, 780, 788, 795, 807 and 811) and threonine (5, 280, 252, 356, 373, 821 and 826) phosphorylation sites are marked schematically with lines (red and green, respectively).

and are not E2F-dependent. Rb is also able to repress transcription from both pol I- and pol III-dependent promoters [23]. Pol I synthesizes large ribosomal RNAs (rRNA) and pol III transcribes genes encoding a variety of small stable RNAs, including 5S rRNA and tRNAs, all of which are essential for cell growth. An attractive hypothesis, therefore, is that by inhibiting pol I and pol III activity Rb may be able to complement its activity as an inhibitor of the chromosome cycle with an ability to put the brakes on cell growth [23].

Structural context

Structural characterisation of cell-cycle proteins to date has focussed on the CDKs. These molecules are regulated by both phosphorylation and non-covalent association with regulatory partners. Crystal structures have illustrated various complexed states of one CDK, CDK2 [24–26]. These structures, together with the structures of the isolated cyclin and CDK components [27,28], suggest that the CDKs are rather plastic molecules, the conformation of which can be altered to affect their activity by the binding of either positive or negative regulatory partners (e.g. cyclin and p27, respectively), as well as by phosphorylation. Only once appropriately phosphorylated and complexed with a cyclin partner, does a CDK adopt its fully catalytically competent conformation. This conformation can be comprehensively disrupted by the binding of a proteinaceous inhibitor molecule, such as a p27. The work of Lee *et al.* [7] extends the range of cell-cycle regulatory processes for which structural information is available downstream from the regulation of CDKs, and into the domain of crucial substrates through which CDKs exert their effect. Here, a structural echo of the CDK regulatory events is observed, as Rb contains two copies of the structural motif found in the cyclin molecule.

The cyclin-box fold connection

As suggested from profile-based sequence analysis [29], and subsequently supported by a threading study [28], the B-pocket region of Rb shares the cyclin-box fold (CBF) that is found in cyclins [28,30,31], transcription factors of the TFIIB family [32] and the A-pocket region

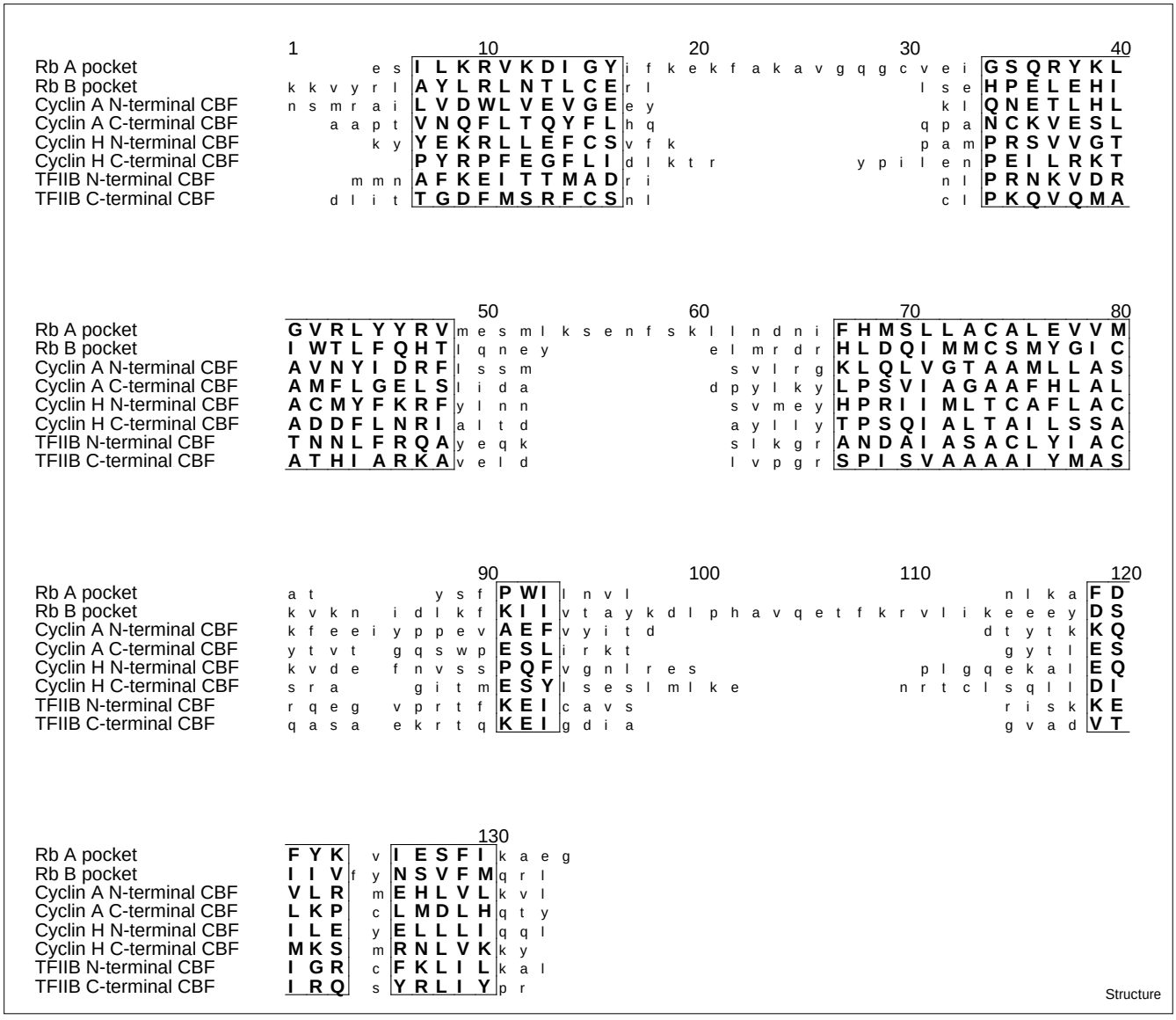
Table 1

Sequence and structural similarity of different cyclin-box folds*.

	Cyclin A		Cyclin H		TFIIB		Rb	
	N	C	N	C	N	C	N	C
Cyclin A								
N	–	2.04	1.64	1.99	1.36	1.52	2.44	2.25
C	11.11	–	1.90	1.83	1.87	1.61	2.04	2.03
Cyclin H								
N	18.81	12.87	–	2.28	1.71	2.27	2.22	2.55
C	9.9	10.89	11.88	–	2.17	2.03	2.39	2.52
TFIIB								
N	14.14	7.45	13.86	6.93	–	1.07	1.89	1.57
C	12.12	10.64	9.9	8.91	20.21	–	1.83	1.65
Rb								
N	8.33	7.41	4.63	6.48	9.26	6.48	–	2.27
C	7.89	6.14	11.4	7.02	14.91	8.77	5.26	–

*The structures of the eight different cyclin-box folds (CBFs) were aligned automatically using the program STAMP, and equivalent residues compared. For each pairwise comparison, this table presents the root mean square deviation for equivalent Cα positions (top right), and the percentage sequence identity for equivalent residues (bottom left).

Figure 2



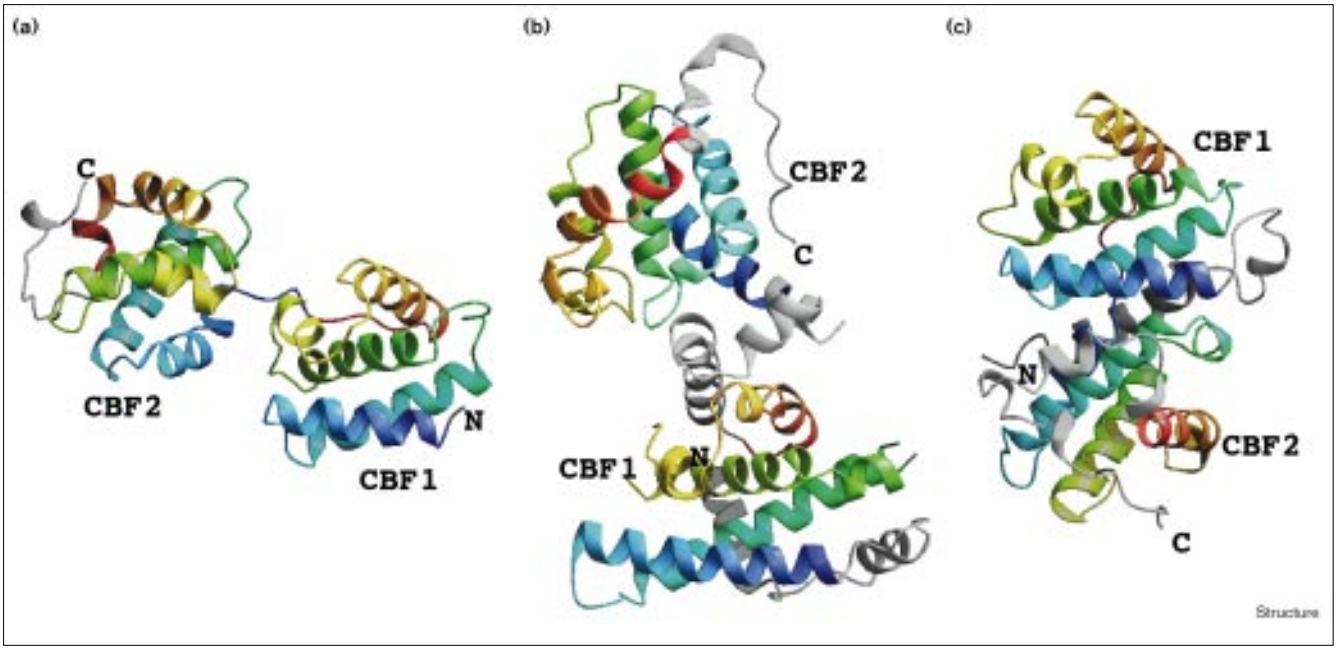
Sequence alignment of cyclin-box fold (CBF) proteins. CBFs were automatically structurally aligned with the program STAMP [43], and the corresponding sequence alignment was illustrated using ALSCRIPT [44]. Boxed regions correspond to areas of structural equivalence.

of Rb [6]. The CBF consists of approximately 100 amino acids which form a bundle of five helices. A structural alignment of the eight available cyclin-box folds (two in each of cyclin A, cyclin H, TFIIB and Rb) reveals sequence identity which varies between five and twenty percent (Table 1). The structural similarity encoded by this sequence identity is surprisingly high, with pairwise comparisons yielding root mean square deviations (rmsds) of between 1.1 Å and 2.5 Å for equivalent Cα positions. Few obvious patterns which might direct adoption of the CBF are apparent from sequence comparisons. Even the previously noted run of four small non-polar amino acids [29] along the buried helix (α3) do not survive the addition of the B-box of Rb to the alignment (Figure 2).

The CBF as a tandemly duplicated motif

Different CBFs are elaborated in different ways by excursions from and extensions to the basic architecture. Confirmation that the B-pocket region of Rb forms a CBF means that in all of the characterised cases, the CBF is found tandemly repeated. Despite this, the relative disposition of the two repeats differs markedly in the different classes of protein (Figure 3), and whether an isolated CBF exists in nature remains open to debate. The strongest

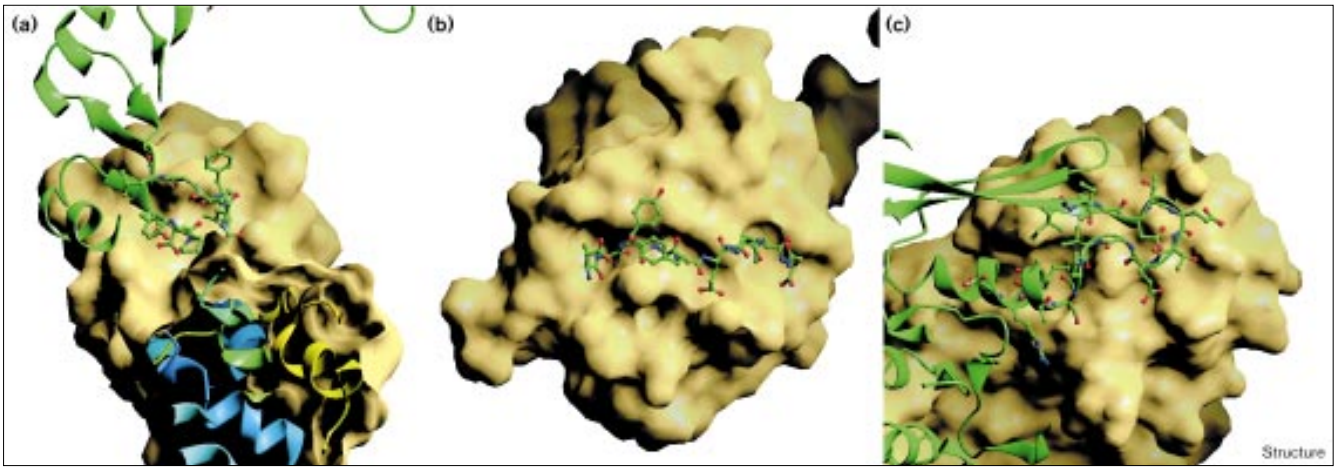
Figure 3



Theme and variation: the CBFs in (a) TFIIB, (b) Rb and (c) cyclin A. In each CBF, residues corresponding to the CBFs are colour-ramped from blue (N-terminal end) to red (C-terminal end). Elaborations of the CBF core are coloured grey. The N and C termini of the proteins are indicated. The N-terminal CBF is marked CBF1 and the C-terminal CBF is marked CBF2. The views are chosen so that the N-terminal CBF of each protein (CBF1) is drawn in the same orientation. From this view, it is apparent that all three proteins have very different relative orientations of the N- and C-terminal CBFs.

suggestion for such an entity comes from the p35 regulatory subunit of CDK5 [33]. No sequence analysis has been able to detect more than a single CBF in the p35 sequence [28], and a 25 kDa fragment of p35 is active in promoting CDK5 activity, despite being scarcely long enough to encode two CBFs [34].

Figure 4



Surfaces of the common interaction site in (a) TFIIB, (b) Rb and (c) cyclin A. The molecular surface of each protein is shown in gold, with the interacting molecule shown in ball-and-stick representation. Each interacting CBF is drawn in the same relative orientation. For Rb, the interacting LXCXE-motif peptide is shown. For cyclin A, the fold of the interacting CDK2 molecule is shown, together with explicit bond representation for residues interacting with this site. For TFIIB, the fold of TBP is shown with explicit bond representation of the interacting residues. TFIIB drawn in this orientation has the C-terminal CBF coming towards the point of view, so that it is clipped by a viewing plane.

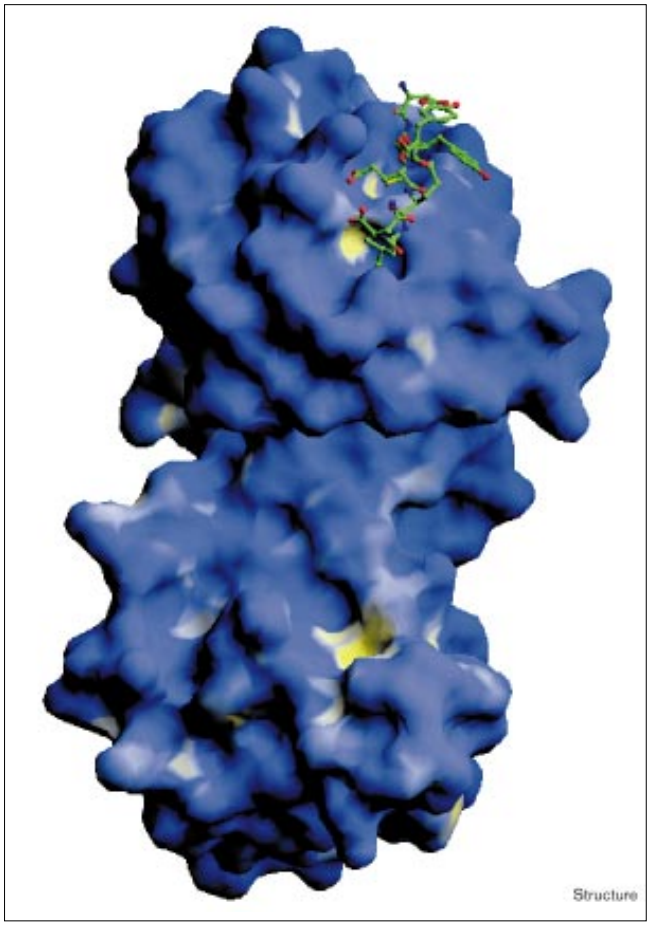
In cyclin A, it is the N-terminal CBF that mediates the so far characterised protein–protein interactions: association with CDK2 and with the inhibitor p27 [24,26]. p27 binds through a conserved sequence motif (RXLFG), found in other members of the p21 CDK-inhibitor (CKI) family, the N-terminal sequences of E2F-1, -2 and -3, and in the p107 linker sequence [35]. In TFIIB, both N-terminal and C-terminal CBFs mediate interactions with DNA and the TATA-binding protein (TBP) [32]. Lee *et al.* have characterised the site of Rb for one important class of interaction: that with proteins which contain the LXCXE motif. A nine-residue peptide derived from the E7 protein of human papilloma virus binds to the B-pocket of Rb. Lee *et al.* postulate that the function of the A-pocket region is to stabilise the fold of the B-pocket region, with which it shares a continuous hydrophobic core. They also note that the A–B interface may be a site of protein–protein interaction.

Interaction sites in CBF proteins

Lee *et al.* observe that sequence conservation and the sites of loss-of-function mutations are concentrated in two regions of Rb: one on the B-pocket region and one at the A/B pocket interface. The former of these two regions is responsible for the binding of the E7 peptide, containing the LXCXE motif. The pattern of alternating conservation and degeneracy across the LXCXE motif is neatly explained by the extended conformation of the bound peptide, so that peptide sidechains point alternately into and away from the binding site. The binding site itself is a groove in the protein surface composed of helices corresponding to the last three structural helices of the B-pocket region CBF. Rb is not alone in using this part of the CBF for protein–protein interactions — it is the same as the shared protein-interaction site identified in TFIIB and cyclin A [36]. In the N-terminal CBF of cyclin A, this groove interacts with the N-terminal loop preceding the PSTAIRE helix of CDK2 [24], whereas in TFIIB, it interacts with the stirrup part of the TBP [32]. Analysis of the surfaces mediating these interactions (Figure 4) shows that Rb has a narrow groove, with occasional deep and hydrophobic wells appropriate to receive residues of the extended LXCXE peptide. In the case of cyclin A, a wider groove accommodates the N-terminal end of the C-helix of CDK2, whereas the preceding loop meanders over the surface of the cyclin molecule. The equivalent groove of TBP is continuous with the cleft separating the N-terminal and C-terminal CBFs, and TPB interacts via residues from both domains. This conserved interaction site contrasts with the p27 RXLFG-recognition site on the N-terminal CBF of cyclin A. In the other CBFs, this site is either inaccessible or unoccupied.

A recently developed protocol (MEMN, unpublished data) for evaluating a hydrophobic potential at a molecular surface has been used to analyse the surface of Rb. This protocol is based on a combination of the GRID algorithm [37] and the Connolly Molecular Surface Package [38]. It has been found in other cases to identify sites of apolar protein–protein interaction, such as the interaction between SH3 and its polyproline ligand, or the LFG motif of p27 with cyclin A. As can be seen in Figure 5, this analysis highlights the hydrophobic pockets of Rb responsible for binding the apolar subsites of the E7 peptide. It also illuminates a patch on the A-pocket region, formed from the interface between α1, α3 and α5. This site is a candidate for some of the additional interactions of Rb, either intermolecular interactions with other proteins or intramolecular interactions with other regions of the Rb protein.

Figure 5



Hydrophobic potential analysis of Rb. The molecular surface of the molecule is coloured according to the GRID hydrophobic potential. A strong preference for a hydrophobic group is indicated by a yellow colour, and no preference for a hydrophobic interaction is indicated by blue colouring. The E7 peptide is shown, interacting with the hydrophobic wells on the B-pocket region (upper). An exposed hydrophobic patch is offered by the A-pocket region (lower).

Outstanding questions

Pocket proteins are not only substrates for members of the CDK family, they can also form stable complexes.

Rb binds to members of the cyclin D family, an interaction that requires the cyclin D LXCXE motif present near its N terminus [39]. The formation of a stable p107–CDK2–cyclin A complex requires the RXLFG sequence motif found in the linker region of p107 [40]. This result suggests that p107 (and E2F-1, -2 and -3) might employ similar structural tactics to those of p27 to bind to certain members of the cyclin family. In addition, the results described by Lee *et al.* show that E2F-1, -2 and -3 can also bind directly to the CBFs of Rb through a second motif, conserved amongst all E2F family members. Binding of p21, and p107 or p130 to CDK2–cyclin complexes, is mutually exclusive [41], although this need not imply similarity in binding mode beyond the RXLFG motif. The crystal structure of a ternary complex formed between CDK2, cyclin A and p27 has revealed that p27 goes on to inhibit CDK2 in a tour de force of molecular disruption, whereas the CDK2 associated with p107 (and E2F-1, -2 or -3) can remain catalytically active [42].

The CBF provides a common structural theme, linking CDK regulation, the control of transcription initiation and the basal transcription machinery. Lee *et al.* have shown details of the interactions between Rb and the LXCXE binding motif, confirming that CBFs use a common set of interaction sites to carry out their functions. The outstanding structural question is how the library of interactions available is used to appropriately assemble and regulate the macromolecular complexes that govern cell-cycle progression.

Acknowledgements

The authors would like to thank Nikola Pavletich for allowing us access to his fascinating structure of the two pocket regions, prior to the PDB release date. We would also like to thank Louise Johnson and Nick Brown for many helpful conversations on how cyclin-box folds might work. Thanks also to Steve Lee for his help in preparing the figures. JAE and MEMN are supported by the Royal Society.

References

1. Friend, S.H., Bernards, R., Rogelj, S., Weinberg, R.A., Rapaport, J.M., Albert, D.M. & Dryja, T.P. (1986). A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. *Nature* **323**, 643-646.
2. Fung, Y.-K.T., Murphree, A.L., T'Ang, A., Qian, J., Hinrichs, S.H. & Benedict, W.F. (1987). Structural evidence for the authenticity of the human retinoblastoma gene. *Science* **236**, 1657-1661.
3. Lee, W.-H., Bookstein, R., Hong, F., Young, L.-J., Shew, J.-Y. & Lee, E.Y.-H.P. (1987). Human retinoblastoma susceptibility gene: cloning, identification and sequence. *Science* **235**, 1394-1399.
4. Horowitz, J.M. *et al.*, & Weinberg, R.A. (1990). Frequent inactivation of the retinoblastoma antioncogene is restricted to a subset of human tumour cells. *Proc. Natl. Acad. Sci. USA* **87**, 2775-2779.
5. Weinberg, R.A. (1995). The retinoblastoma protein and cell cycle control. *Cell* **81**, 323-330.
6. Kim, H. & Cho, Y. (1997). Structural similarity between the pocket region of retinoblastoma tumour suppressor and the cyclin-box. *Nat. Struct. Biol.* **4**, 390-395.
7. Lee, J.-O., Russo, A.A. & Pavletich, N.P. (1998). Structure of the retinoblastoma tumour-suppressor pocket domain bound to a peptide from HPV E7. *Nature* **391**, 859-865.
8. Johnson, D.G., Schwarz, J.K., Cress, W.D. & Nevins, J.R. (1993). Expression of transcription factor E2F1 induces quiescent cells to enter S phase. *Nature* **365**, 349-352.
9. Weintraub, S.J., Prater, C.A. & Dean, D.C. (1992). Retinoblastoma protein switches the E2F site from positive to negative element. *Nature* **358**, 259-261.
10. Muller, R. (1995). Transcriptional regulation during the mammalian cell cycle. *Trends Genetics* **11**, 173-178.
11. Helin, K. (1998). Regulation of cell proliferation by the E2F transcription factors. *Curr. Opin. Genetics Dev.* **8**, 28-35.
12. Lees, J.A., Buchkovich, K.J., Marshak, D.R., Anderson, C.W. & Harlow, E. (1991). The retinoblastoma protein is phosphorylated on multiple sites by human cdc2. *EMBO J.* **10**, 4279-4290.
13. Mittnacht, S., Lees, J.A., Desai, D., Harlow, E., Morgan, D.O. & Weinberg, R.A. (1994). Distinct sub-populations of the retinoblastoma protein show a distinct pattern of phosphorylation. *EMBO J.* **13**, 118-127.
14. Taya, Y. (1997). RB kinases and RB-binding proteins: new points of view. *Trends Biochem. Sci.* **22**, 14-17.
15. Kitagawa, M., *et al.*, & Taya, Y. (1996). The consensus motif for phosphorylation by cyclin D1-Cdk4 is different from that for phosphorylation by cyclin A/E-Cdk2. *EMBO J.* **15**, 7060-7069.
16. Zarkowska, T. & Mittnacht, S. (1997). Differential phosphorylation of the retinoblastoma protein by G1/S cyclin-dependent kinases. *J. Biol. Chem.* **272**, 12738-12746.
17. Knudsen, E.S. & Wang, J.Y.J. (1996). Differential regulation of retinoblastoma protein function by specific CDK phosphorylation sites. *J. Biol. Chem.* **271**, 8313-8320.
18. Mittnacht, S. (1998). Control of pRB phosphorylation. *Curr. Opin. Genetics Dev.* **8**, 21-27.
19. Wang, J.Y.J. (1997). Retinoblastoma protein in growth suppression and death protection. *Curr. Opin. Gen. Dev.* **7**, 39-45.
20. Hu, Q., Dyson, N. & Harlow, E. (1990). The regions of the retinoblastoma protein needed for binding to adenovirus E1A or SV40 large T antigen are common sites for mutations. *EMBO J.* **9**, 1147-1155.
21. Welch, P.J. & Wang, J.Y.J. (1993). A C-terminal protein-binding domain in the retinoblastoma protein regulates nuclear c-Abl tyrosine kinase in the cell cycle. *Cell* **75**, 779-790.
22. Xiao, Z.-X., *et al.,* & Livingston, D.M. (1995). Interaction between the retinoblastoma protein and the oncoprotein MDM2. *Nature* **375**, 694-698.
23. White, R.J. (1997). Regulation of RNA polymerases I and III by the retinoblastoma protein: a mechanism for growth control? *Trends Biochem. Sci.* **22**, 77-80.
24. Jeffrey, P. & Pavletich, N. (1995). Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. *Nature* **376**, 313-320.
25. Russo, A.A., Jeffrey, P.D. & Pavletich, N.P. (1996). Structural basis of cyclin-dependent kinase activation by phosphorylation. *Nat. Struct. Biol.* **3**, 696-700.
26. Russo, A.A., Jeffrey, P.D., Patten, A.K., Massague, J. & Pavletich, N.P. (1996). Crystal structure of the p27kip1 cyclin-dependent-kinase inhibitor bound to the cyclinA–Cdk2 complex. *Nature* **382**, 325-331.
27. DeBondt, H.L., Rosenblatt, J., Jancarik, J., Jones, H.D., Morgan, D.O. & Kim, S.-H. (1993). Crystal structure of cyclin-dependent kinase 2. *Nature* **363**, 595-602.
28. Brown, N.R., *et al.*, & Johnson, L.N. (1995). The crystal structure of cyclin A. *Structure* **3**, 1235-1247.
29. Gibson, T.J., Thompson, J.D., Blocker, A. & Kouzarides, T. (1994). Evidence for a protein domain superfamily shared by the cyclins, TFIIB and RB/p107. *Nucleic Acids Res.* **22**, 946-952.
30. Andersen, G., Poterszman, A., Egly, J., Moras, D. & Thierry, J. (1996). The crystal structure of human cyclin H. *FEBS Lett.* **397**, 65-69.
31. Kim, K., Chamberlin, H., Morgan, D. & Kim, S. (1996). Three-dimensional structure of human cyclin H, a positive regulator of the CDK-activating kinase. *Nat. Struct. Biol.* **3**, 849-55.
32. Nikolov, D.B., *et al.* & Burley, S.K. (1995). Crystal structure of a TFIIB-TBP-TATA-element ternary complex. *Nature* **377**, 119-128.
33. Tsai, L.-H., Delalle, I., Caviness Jr, V.S., Chae, T. & Harlow, E. (1994). p35 is a neural-specific regulatory subunit of cyclin-dependent kinase 5. *Nature* **371**, 419-423.
34. Lew, J., *et al.*, & Wang, J.H. (1994). A brain-specific activator of cyclin-dependent kinase 5. *Nature* **371**, 423-426.

35. Adams, P., Sellers, W., Sharma, S., Wu, A., Nalin, C. & Kaelin WG, J. (1996). Identification of a cyclin-cdk2 recognition motif present in substrates and p21-like cyclin-dependent kinase inhibitors. *Mol. Cell Biol.* 16, 6623-33.
36. Noble, M.E.M., Endicott, J.A., Brown, N.R. & Johnson, L.N. (1997). The cyclin-box fold: protein recognition in cell-cycle and transcription control. *Trends Biochem. Sci.* 22, 482-487.
37. Goodford, P. (1996). Multivariate characterization of molecules for QSAR. *J. Chemometrics* 10, 107-117.
38. Connolly, M. (1985). Molecular surface triangulation. *J. Appl. Cryst.* 18, 499-505.
39. Dowdy, S., Hinds, P., Louie, K., Reed, S., Arnold, A. & Weinberg, R. (1993). Physical interaction of the retinoblastoma protein with human D cyclins. *Cell* 73, 499-511.
40. Faha, B., Ewen, M., Tsai, L., Livingston, D. & Harlow, E. (1992). Interaction between human cyclin A and adenovirus E1A-associated p107 protein. *Science* 255, 87-90.
41. Zhu, L., Harlow, E. & Dynlacht, B. (1995). p107 uses a p21CIP1-related domain to bind cyclin/cdk2 and regulate interactions with E2F. *Genes Dev.* 9, 1740-52.
42. Hauser, P.J., Agarwal, D., Chu, B. & Pledger, W.J. (1997). p107 and p130 associated cyclin A has altered substrate specificity. *J. Biol. Chem.* 272, 22954-22959.
43. Russell, R.B. & Barton, G.J. (1992). Multiple sequence alignment from tertiary structure comparison: assignment of global and residue confidence levels. *Proteins*. 14, 309-323.
44. Barton, G.J. (1993). ALSCRIPT: A tool to format multiple sequence alignments. *Protein Eng.* 6, 37-40.